UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/670,251	09/26/2000	Mukund Padmanabhan	YOR92000390	5892

35195   7590   08/29/2005

FERENCE & ASSOCIATES
409 BROAD STREET
PITTSBURGH, PA 15143

EXAMINER
LERNER, MARTIN

ART UNIT	PAPER NUMBER
2654	

DATE MAILED: 08/29/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) FROM
THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed
  after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any
  earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>28 July 2005</u>.

2a)☒ This action is **FINAL**.   2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is
closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 to 15* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1 to 15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All   b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage
application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☐ Notice of References Cited (PTO-892)
2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3)☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
    Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.
5)☐ Notice of Informal Patent Application (PTO-152)
6)☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 102

1.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

2.   Claims 1, 7, 8, 14, and 15 are rejected under 35 U.S.C. 102(b) as being

anticipated by *Woodland et al. ("Iterative Unsupervised Adaptation Using Maximum*

Likelihood Linear Regression").

Regarding independent claims 1, 8, and 15, *Woodland et al.* discloses a method,

apparatus, and computer program for adaptation in speech recognition, comprising:

"providing at least one speech recognition model" – gender independent Hidden

Markov Models (HMMs) HMM-1 and HMM-2 are built from acoustic training data sets

consisting of 36,493 sentences (Page 1134, Left Column, Paragraphs 4 to 7; Page

1135, Right Column, Paragraph 3);

"accepting speaker data" – test H3-P0 data was captured for each speaker of 20

speakers (Page 1134, Right Column, Paragraph 7);

"generating a word lattice having a plurality of paths based on the speaker data"

– H3 development test data is used for lattice generation (Page 1135, Left Column,

Paragraphs 3 to 5: Table 1); word lattices are used to generate an error rate for H3-P0

data (Page 1136, Left Column, Lines 1 to 6: Table 3); implicitly, a word lattice has a

plurality of paths;

"wherein the generation of the word lattice takes into account language model

probabilities" – the HTK LVCSR system uses a decoder to produce word lattices

containing language model information for the application for rescoring of new language

models (Page 1134, Left Column, Paragraph 8); lattices generated by the HTK system

contain a set of nodes that correspond to particular instants and arcs connecting these

nodes that represent hypotheses for the time period between the two nodes; associated

with each arc are both language model and acoustic model scores; lattices may contain

copies of each word, and further copies can be required to encode the language model

constraints (Page 1135. Left Column, Paragraph 2); language model scores correspond

to "language model probabilities";

"adapting at least one of the speaker data and the at least one speech

recognition model with respect to the generated word lattice in a manner to maximize

the likelihood of the speaker data" – language models were trained on the text training

corpus and the H3 text data sets; HMM-1 models used global MLLR adaptation and

specific MLLR adaptation from word lattices for H3-P0 data; the result is a decreased

error rate by adapting HMM-1 ("speech recognition model") to H3 data ("speaker data")

using MLLR (Maximum Likelihood Linear Regression") (Page 1135, Right Column,

Paragraph 5 to Page 1136, Right Column, Paragraph 2: Table 3).

Regarding claims 7 and 14, *Woodland et al.* discloses maximum likelihood linear

regression (MLLR) for adaptation of speaker data in speech recognition (Page 1133).

Claim Rejections - 35 USC § 103

3.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4.   Claims 2 to 6 and 9 to 13 are rejected under 35 U.S.C. 103(a) as being

unpatentable over *Woodland et al.* in view of *Nguyen et al.*

Concerning claims 2 and 9, *Woodland et al.* discloses generating word lattices,

but omits generating word lattices by maximum a-posteriori adaptation. However,

Nguyen et al. teaches adaptation by both Maximum Likelihood Linear Regression

(MLLR) and Maximum A Posteriori (MAP) adaptation, noting that both techniques are

available to perform adaptation. It is stated that Bayesian-based MAP techniques are

particularly useful in dealing with adaptation of sparse data sets, but in practical

applications, depending upon the amount of adaptation data available, a combination of

both MLLR and MAP may be used. (Column 1, Lines 50 to 60) Thus, *Nguyen et al.*

performs adaptation with both MLLR and MAP. (Column 3, Line 57 to Column 4, Line

32) It would have been obvious to one having ordinary skill in the art to generate a

word lattice with maximum a posteriori adaptation as taught by *Nguyen et al.* in MLLR

adaptation with word lattices of *Woodland et al.* for the purpose of dealing with

adaptation of sparse data sets in H3 training data.

Concerning claims 3 and 10, *Nguyen et al.* discloses Bayesian adaptation by

MAP with Equation 4; γ is the observed posterior probability of the observation to adapt

the speech models ("posterior state occupancy probability"); μ_{MAP} is found by summing

the observed posterior probabilities over time: $\Sigma\, \gamma(t)\, o_t$ and $\Sigma\, \gamma(t)$ ("posterior word

occupancy probabilities by summing over all states interior to a word") (column 4, lines

10 to 23); the adaptation system then processes the segments in an N-best pass to

collect the most probable labels; model adaptation may be performed to adapt speech

models to words ("at least one likely word at each frame") (column 3, lines 1 to 8;

column 3, lines 47 to 56).

Concerning claims 4 and 11, *Woodland et al.* discloses word lattices (Page 1135,

Left Column); a word lattice implicitly contains word traces.

Concerning claims 5 and 12, *Woodland et al.* discloses pruning during adaptation

(Paragraph Bridging Pages 1135 to 1136), but does not expressly discard

interpretations associated with low confidence. However, *Nguyen et al.* teaches

assigning weights to the N-best transcriptions, so that reliable information becomes

enhanced by a positive weight, and unreliable information is correspondingly diminished

by a negative weight. The system thus tends to push models that generate incorrect

labels away from those that generate correct ones. Subsequently, model information is

accumulated among the N-best transcriptions for the entire set of sentences and then

used to adapt the speech models. (Column 3, Lines 32 to 56; Column 4, Lines 23 to

59) Taking the N-best of the most reliable transcriptions necessarily implies eliminating

transcriptions not associated with the N-best most reliable transcriptions ("discarding

interpretations associated with low confidence"). N-best techniques are well known in speech recognition. *Nguyen et al.* says assigning weights to the N-best transcriptions corresponding to their likelihoods produces a natural information and data corrective process. (Column 3, Lines 31 to 34) It would have been obvious to one having ordinary skill in the art to utilize the N-best technique of *Nguyen et al.* to discard unreliable transcriptions for pruning in MLLR adaptation with word lattices of *Woodland et al.* for the purpose of producing a natural information corrective process.

Concerning claims 6 and 13, *Nguyen et al.* discloses Bayesian adaptation by MAP with Equation 4; γ is the observed posterior probability of the observation to adapt the speech models ("posterior phone probability") (column 4, lines 10 to 23); the observations and labels represent phonemes in speech recognition.

Response to Arguments

5.   Applicants' arguments filed 28 June 2005 have been fully considered but they are not persuasive.

Applicants argue that *Woodland et al.* omits the limitation of "generating a word lattice having a plurality of paths based on the speaker data, wherein the generation of a word lattice takes into account language model probabilities." This is not persuasive.

The Advisory Action mailed 13 July 2005 stated that *Woodland et al.* clearly discloses a word lattice taking into account language model probabilities. *Woodland et al.* says the HTK LVCSR system uses a decoder to produce word lattices containing language model information for the application for rescoring of new language models.

(Page 1134, Left Column, Paragraph 8) Lattices generated by the HTK system contain a set of nodes that correspond to particular instants and arcs connecting these nodes that represent hypotheses for the time period between the two nodes. Associated with each arc are both language model and acoustic model scores. Lattices may contain copies of each word, and further copies can be required to encode the language model constraints. (Page 1135, Left Column, Paragraph 2) Language model scores correspond to "language model probabilities". Indeed, Applicants note *Woodland et al.* discloses global MLLR adaptation with trigram language models and bigram language models. (Page 1135, Right Column, Last Paragraph to Page 1136, Left Column, First Paragraph) Thus, *Woodland et al.* clearly discloses the limitation of "generating a word lattice having a plurality of paths based on the speaker data, wherein the generation of a word lattice takes into account language model probabilities."

Also, Applicants maintain that they improve the performance of MLLR, even when the unsupervised transcription is mildly erroneous, by taking into account the fact that the initial transcription contains errors. Applicants say they accomplish this by considering not just the "1-best" (*i.e.* the single best) transcription produced during the first pass decoding, but the top N candidates. Applicants state that there is no teaching or suggestion of using more than one transcription during adaptation by *Woodland et al.* This position is not persuasive.

Firstly, Applicants' claims do not expressly include any limitation directed to transcriptions produced by the top N candidates. Only Applicants' Specification discloses the cited feature. Although the claims are interpreted in light of the

specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Secondly, *Nguyen et al.* clearly discloses the feature of adaptation with N-best transcriptions. Thus, even if the limitation were expressly claimed, the feature would at least be obvious over *Woodland et al.* in view of *Nguyen et al.*

Moreover, it is noted in the Advisory Action mailed 13 July 2005 that *Woodland et al.* clearly disclosed the cited feature of a word lattice taking into account language model probabilities, and indicated that no new issues were involved requiring further consideration and/or search, and no issues of new matter were presented. Thus, a First Action Final Rejection is proper, as the same references and same grounds of rejection are applied as an immediately prior final rejection.

Therefore, the rejections of claims 1, 7, 8, 14, and 15 under 35 U.S.C. 102(b) as being anticipated by *Woodland et al.*, and of claims 2 to 6 and 9 to 13 under 35 U.S.C. 103(a) as being unpatentable over *Woodland et al.* in view of *Nguyen et al.*, are proper.

Conclusion

6.   This is a Request for Continued Examination (RCE) of Applicants' earlier Application No. 09/670,251. All claims are drawn to the same invention claimed in the earlier application and could have been finally rejected on the grounds and art of record in the next Office action if they had been entered in the earlier application. Accordingly, **THIS ACTION IS MADE FINAL** even though it is a first action in this case. See MPEP

§ 706.07(b). Applicants are reminded of the extension of time policy as set forth in 37
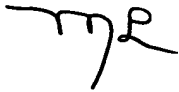
CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no, however, event will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Martin Lerner whose telephone number is (571) 272-

7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to
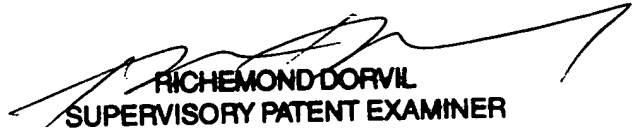
Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone

number for the organization where this application or proceeding is assigned is 703-

872-9306.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free).

ML
8/23/05

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER